DSpace Evaluation

Jonathan Lorig

Metadata Review Group Meeting

NASA Goddard Space Flight Center Library

March 10, 2005

Agenda

- Presenter background
- Institutional Repository requirements for DAS
- DSpace
  - Background
  - Findings
  - Next steps

Presenter Background

- Digital Librarian at GSFC; started Jan ‘05
- Coordinate with Gail and Nikkia to research and develop digital mechanisms for knowledge preservation at Goddard
- Worked last summer at Glenn with Sue Oberc; developed a serials management dataset and fought with the Sirsi Workflows report module
- BS in Civil Engineering, MS in Library Science from University of Illinois

Requirements for the DAS Institutional Repository

- Capture, catalog, preserve, and redistribute GSFC intellectual content in various digital formats
- Provide single search interface
- Utilize the Goddard Core Metadata Element Set
- Provide workflow authority settings for metadata and object ingest
- Provide separate metadata and object access security
- Provide access information for objects stored on secure servers outside the Goddard Library

Current wishes for the DAS

- Provide customizable metadata schema for individual GSFC projects
- Integrate front end for website spider captures and automated metadata extraction
- Customize persistent identifiers and provide secure handle resolution

DSpace

- IR platform written by programmers at MIT and Hewlett-Packard labs
- Open-source Java permits customization and collaboration with other institutions
- Implemented so far by 68 universities and technical libraries according to DSpace website
- Langley and JPL librarians will present on their DSpace implementations at an STI ViTS session tomorrow

DSpace User Community offers Substantial Contributions

- Ongoing core code development by original developers
- SIMILE - MIT project to extend metadata schema support through RDF and semantic web technologies
- DSpace/ARK - UCLA/UCSD/CDL project to support alternative identifier schemes
- XML/XSLT - TAMU project to replace HTML code with XML code throughout DSpace interface

DSpace Projects Continued….

- DRUM - Digital Repository at the University of Maryland allows faculty to authenticate against campus LDAP directory with secure user profiles
- Research Tools - University of Rochester project permits customized project information pages
- OAI Harverster Plug-In from Old Dominion University Digital Library Research Group

DSpace testing and external reports so far have indicated….

- Robust ingest workflow settings including customizable workflows for collections
- Customizable user interface
- Customizable global metadata scheme; still evaluating schema for individual collections
- Simplistic metadata and object access controls would require substantial administrative effort
- Substantial community support and quick answers to technical questions

DSpace code uses languages and standards identical to the DAS

- Java servlets and Java Server Pages (JSP) for base code and page display
- Lucene search engine for metadata indexing and retrieval
- OAI-PMH compliance
- SQL database compatibility
- DSpace also supports OpenURL protocol from SFX for potential interoperability with external library databases

Next Steps to Evaluate DSpace

- Complete test of customizable metadata schema using current DAS content
- Identify preservation issues pertaining to DSpace's bitstream method for object storage
- Identify Java code for web capture and metadata extraction to append to DSpace (Heritrix?)
- Identify code for additional administrative tools to control object security and distribution settings
  - specific privileges for particular users in sub-communities and collections
  - options for various object display formats or storage of multiple formats